Hospital efficiency analysis through individual effects

Koop, G.; Osiewalski, J.; Steel, M.F.J.

Publication date:
1994

CentER
for
Economic Research

Discussion paper

TILBURG UNIVERSITY

44

Center
for
Economic Research

No. 9447

HOSPITAL EFFICIENCY ANALYSIS THROUGH INDIVIDUAL EFFECTS: A BAYESIAN APPROACH

by Gary Koop, Jacek Osiewalski,
and Mark F.J. Steel

R11

June 1994

ISSN 0924-7815

Hospital Efficiency Analysis Through Individual Effects:

A Bayesian Approach

Gary Koop
Department of Economics
University of Toronto

Jacek Osiewalski
Department of Econometrics
Academy of Economics, Kraków

Mark F. J. Steel
CentER and Department of Econometrics
Tilburg University

ABSTRACT: This paper develops Bayesian tools for making inferences about firm-specific inefficiencies in panel data models. We begin by establishing a Bayesian setting in which fixed and random effects models are defined. What distinguishes both classes of models is the marginal prior independence of the effects. These techniques are applied to a panel of U.S. hospitals. Our empirical findings illustrate the different characteristics of both types of models, as well as the influence of the particular priors used on the firm effects. In addition, we find that variables such as non-profit, for-profit or government-run dummies and an average Herfindahl index have little explanatory power for hospital efficiencies.

KEYWORDS: stochastic frontiers; panel data; fixed effects; random effects; Monte Carlo Markov chains

JEL classification: C11, C23, I10

ACKNOWLEDGEMENTS: We would like to thank the VA Management Science Group, Bedford, Mass., for providing the data used in this study. We benefited from helpful comments by Dale Poirier and Michael Denny. The second author acknowledges the hospitality of the Department of Econometrics, Tilburg University.

CORRESPONDENCE TO: Mark F.J. Steel, Department of Econometrics, Tilburg

University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.
Phone: +31-13-66 3246; Fax: +31-13-66 3280; E-mail: steel@kub.nl

1. Introduction

Cost containment in health care is one of the most important public policy issues in modern America. Since hospital spending is a large part of total health care spending, an understanding of the cost structure of hospitals is crucial. This paper uses data from a large panel of U.S. hospitals from 1987-1991 to investigate hospital efficiencies. We use the stochastic cost frontier methodology for panel data described in, for instance, Schmidt and Sickles (1984). We introduce firm-specific or individual effects, which are assumed to be constant over time. No time effects will be considered in this paper. This methodology is typically implemented using one of two approaches which are often called the fixed and random effects models. The differences between these two types of models can be viewed, from a Bayesian perspective, as a difference in the structure of the prior information.

We develop a Bayesian framework in which we define fixed effects and random effects models. Our Bayesian fixed effects models are characterized by marginal prior independence between the individual effects, which are thus not linked across firms, but are only assumed constant over time. We distinguish the **standard individual effects (SIE) model**, where an improper prior on individual intercepts is used. We call our second fixed effects model the **marginally independent efficiency distribution (MIED) model**. In the latter model we use a proper prior on the firm-specific effects, which are still independent. In this context, the stochastic frontier interpretation implies that the inefficiency error should be one-sided, and the Bayesian approach allows us to easily incorporate this prior information.

In the so-called Bayesian random effects models we assume prior links between the individual effects: their means can be functionally related to certain firm characteristics, which defines the **varying efficiency distribution (VED) model**, or they can all be drawn from a common distribution, leading to the **common efficiency distribution (CED) model**.

With respect to the existing Bayesian literature, i.e. van den Broeck, Koop, Osiewalski and Steel (1994) and Koop, Osiewalski and Steel (1994), we have incorporated the following methodological advances: i) treatment of panel data as opposed to a cross-section analysis, ii) explicitly allowing efficiencies to depend on firm characteristics, iii) providing Bayesian counterparts to the classical fixed and random effects models.

Our application indicates that the techniques we propose are computationally feasible and yield sensible results. We explicitly state the consequences of certain prior assumptions, and illustrate the differences between Bayesian fixed and random effects models. The estimated frontier is largely consistent with economic regularity conditions, and prior sensitivity of inference on efficiencies is examined. However, we find that the types of variables which we expect to affect efficiency (i.e. non-profit, for-profit or government-run dummies and a Herfindahl index), do not do so in a systematic manner.

2. Efficiency Analysis with Panel Data

In order to measure hospital efficiencies, we adopt the stochastic frontier framework first developed by Meeusen and van den Broeck (1977) and Aigner, Lovell and Schmidt (1977). This methodology postulates a cost frontier, reflecting technology common to all firms, which represents the minimum attainable cost of producing a given level of output(s). Deviations from this frontier reflect either measurement error or inefficiency. Measurement error is symmetric and Normally distributed while inefficiency has some one-sided distribution. For instance, if y_i is the log of costs of firm i and x_i is a vector of k appropriate explanatory variables, then a typical stochastic frontier model may be specified as:

$$y_i=\alpha_0+x_i'\beta+v_i+u_i, \tag{1}$$

where the k-dimensional vector β describes the frontier, u_i and v_i are independent of each other, v_i is i.i.d. Normal and u_i is i.i.d. with some one-sided distribution, i=1,..,N. Given these assumptions, the likelihood function can be derived and inferences made about the firm specific inefficiencies. We analyze overall productive efficiency in the sense of Farrell (1957) (see Kopp and Diewert (1982)). We interpret (1) as defining a conditional model for y_i given x_i, and thus we assume that sufficient conditions for the weak exogeneity of x_i are fulfilled (see Engle, Hendry and Richard (1983)). On account of the independence across firms, we can then validly use (1) for predictions corresponding to, as yet, unobserved firms, as explained in Osiewalski and Steel (1994).

In previous work (van den Broeck, Koop, Osiewalski and Steel (1994)), we have argued for the adoption of a Bayesian perspective for making inferences from such models, since such an approach yields exact finite sample results, allows us to mix over models, to conduct inference on the actual efficiencies, and surmounts some difficult statistical issues which arise in classical analyses. In the present application, we have panel data and new issues arise which necessitate an extension of our previous work. The classical econometric analysis of firm efficiency with panel data is described in Schmidt and Sickles (1984). If we extend (1) to allow for a time component (t=1,..,T) and assume that efficiency is constant over time for a given firm, we obtain:

$$y_i=(\alpha_0+u_i)\iota_T+X_i\beta+v_i, \tag{2}$$

where y_i is now a Tx1 vector containing observations for firm i, X_i is Txk, ι_T is a Tx1 vector of ones, and v_i is i.i.d. $N(0,\sigma^2 I_T)$, i = 1,...,N. There are two ways of proceeding with this model, which correspond to the distinction between the fixed effects and the random effects model used in panel data analysis.

In the classical fixed effects model, we define $\alpha_i=\alpha_0+u_i$, where the u_i's are the individual effects. The inefficiency is thus associated with the firm-specific intercept. Schmidt and Sickles define $\hat{u}_i=\hat{\alpha}_i-\hat{\alpha}$, where $\hat{\alpha}_i$ are ordinary least squares estimates of the intercepts and $\hat{\alpha}=\min_j(\hat{\alpha}_j)$. The $\hat{u}_i$'s (or $\exp(-\hat{u}_i)$'s) are used as measures of inefficiency (efficiency). Note that this approach assumes that the firm with the smallest $\hat{\alpha}_i$ is fully efficient and measures inefficiencies as deviations from this firm, i.e. it leads to the analysis of relative efficiencies. Furthermore, the use of the min operator above makes the classical distribution theory for the $\hat{u}_i$'s difficult and, hence, it is hard to calculate standard errors for the efficiencies. The Bayesian approach provides the tools to surmount such problems, as will be stressed below.

In the classical random effects model, it is assumed that u_i has some one-sided distribution and maximum likelihood estimation can be carried out as described in Pitt and Lee (1981) and Schmidt and Sickles (1984). The disadvantage of this approach over the fixed effects approach is that a distributional assumption must be made for the inefficiencies. The advantage is that the problems with the distributional theory of the $\hat{u}_i$'s described above do not occur. Furthermore, Simar (1992) argues that, in practice, the fixed effects model may produce poor estimates of the parameters and the efficiencies in the (usual) case where $T<<N$ and regressors do not vary much over time.

McCulloch and Rossi (1994) rightly remark that "in the Bayesian point of view, there is no distinction between fixed and random effects models, only between hierarchical and non-hierarchical models". However, we feel it is useful to construct a Bayesian framework in which the widely used panel data terminology is preserved, and fixed and random effects models are clearly distinguished. The distinguishing characteristic of both groups of models will, of course, not be the deterministic or random nature of the effects, but rather the marginal prior links between the effects. Typically, these N effects will be assumed to be prior independent, either marginally, which gives us the Bayesian fixed effects structure, or conditionally upon a small number $m<<N$ of additional parameters, which link the effects and introduce marginal dependence of these effects across time or individuals. In the latter case, we talk of Bayesian random effects models, which, by construction, require non-trivial hierarchical prior structures.

2.1 Bayesian Fixed Effects Models

Standard Individual Effects (SIE) Model

We start from the basic individual effects model in (2). As in the classical analysis, we define $\alpha_i=\alpha_0+u_i$ and let $\alpha=(\alpha_1...\alpha_N)'$. Under the standard noninformative prior, $p(\alpha,\beta,\sigma^{-2})\propto\sigma^2$, the full Bayesian model, therefore, is given by:

$$p(y,\alpha,\beta,\sigma^{-2}\mid X)=c\sigma^2\prod_{i=1}^{N} f_N^T(y_i\mid X_i\beta+\alpha_i\iota_T,\sigma^2 I_N) \tag{3}$$

where $c>0$, and $f_N^T(.\mid a,B)$ denotes the probability density function of a T-variate Normal distribution with mean vector a and covariance matrix B. We immediately infer from (3) that the marginal posterior distribution of (α,β) is the $(N+k)$-variate Student-t distribution with $N(T-1)-k$ degrees of freedom. In our application $N=382$ and $T=5$, so degrees of freedom are around 1,500 (the exact value depends on k). In the light of this, the Student-t posterior will be almost identical to a Normal distribution and, for the rest of the SIE case, we present results in terms of this Normal approximation.

The marginal posterior for the parameters of the cost frontier, β, is the kvariate Normal distribution with

$$E(\beta \mid y,X)=\hat{\beta}=S^{-1}\sum_{i=1}^{N}(X_i-\iota_T\bar{x}_i')'(y_i-\bar{y}_i\iota_T), \tag{4}$$

where

$$\bar{x}_i=\frac{1}{T}X_i'\iota_T,$$
$$\bar{y}_i=\frac{1}{T}\iota_T'y_i,$$
$$S_i=(X_i-\iota_T\bar{x}_i')'(X_i-\iota_T\bar{x}_i'),$$

and

$$S=\sum_{i=1}^{N}S_i.$$

Equation (4) is the so-called "within estimator" from the panel data literature. The covariance matrix for the marginal posterior for β is given by

$$V(\beta \mid y,X)=\hat{\sigma}^2S^{-1}, \tag{5}$$

where

$$\hat{\sigma}^2=\frac{1}{N(T-1)-k}\sum_{i=1}^{N}(y_i-\hat{\alpha}_i\iota_T-X_i\hat{\beta})'(y_i-\hat{\alpha}_i\iota_T-X_i\hat{\beta}).$$

The marginal posterior of α is the N-variate Normal distribution with means

$$E(\alpha_i \mid y,X)=\hat{\alpha}_i=\bar{y}_i-\bar{x}_i'\hat{\beta} \qquad i=1,..,N, \tag{6}$$

and covariances

$$cov(\alpha_i,\alpha_j \mid y,X)=\hat{\sigma}^2\left(\frac{\delta(i,j)}{T}+\bar{x}_i'S^{-1}\bar{x}_j\right) \qquad i,j=1,..,N, \tag{7}$$

where $\delta(i,j)=1$ if $i=j$ and 0 otherwise.

In the present paper, interest centres on firm-specific and predictive efficiency. In the classical analysis of Schmidt and Sickles (1984), the authors assumed one firm was fully efficient and measured inefficiencies relative to this firm. In our Bayesian analysis of this model with an improper uniform prior on α, we, too, need to measure inefficiency relative to the most efficient firm, but we do not assign this status to one particular firm. In fact, our approach allows us to formally treat the uncertainty implicit in deciding which is the most efficient firm. That is, due to parameter uncertainty, it is not necessarily the case that the firm with the smallest $\hat{\alpha}_i$ is the most efficient. Formally, we define relative firm-specific efficiency as $r_i^{rel}=\exp(-u_i^{rel})$, where $u_i^{rel}=u_i-\min_j(u_j)=\alpha_i-\min_j(\alpha_j)$. The nonnegativity restriction, following directly from the interpretation of the inefficiency term, is not imposed on u_i here, but rather on u_i^{rel}. As the definition of u_i^{rel} depends on the number of firms in the sample under consideration, this makes the implicit prior on r_i^{rel} a function of N. The implied prior distributions for the efficiencies, r_i^{rel}, are characterized by a point mass of 1/N at full efficiency ($r_i^{rel}=1$) and $p(r_i^{rel}) \propto 1/r_i^{rel}$ for $r_i^{rel}\in(0,1)$. The latter is an L-shaped improper density, which for an arbitrarily small a ϵ (0,1) puts an infinite mass in (0,a), but only a finite mass in [a,1). Thus the implied prior strongly favors low efficiency.

If we knew which firm was most efficient, it would be straightforward to calculate inefficiencies relative to this firm. However, we do not know this, and must formally incorporate this uncertainty into the analysis. The marginal posterior distributions of u_i^{rel} and of the relative efficiency of the i'th firm, r_i^{rel}, have a point mass at full efficiency given by:

$$P(u_i^{rel}=0\mid y,X)=P(\alpha_i=\min_j\alpha_j\mid y,X)\equiv P_i, \tag{8}$$

and we have the following density function for $u_i^{rel}>0$:

$$\begin{aligned} p(u_i^{rel}\mid y,X) &= \sum_{j=1,j\neq i}^{N} p(u_i^{rel}\mid y,X,u_j^{rel}=0)P(u_j^{rel}=0\mid y,X) \\ &= \sum_{j=1,j\neq i}^{N} P_j p(u_i^{rel}\mid y,X,u_j^{rel}=0). \end{aligned} \tag{9}$$

In other words, we calculate the distribution of the efficiency of firm i relative to firm j for all j, and then weight by the probability that the j'th firm is the most efficient. This allows us to calculate exact, finite sample results for the relative efficiencies, and deals with the difficult distributional issues which arise in the classical analysis of Schmidt and Sickles as a result of the min operator being present.

It remains to describe how to calculate P_j and $p(u_i^{rel}|y,X,u_j^{rel}=0)$. P_j is the probability that firm j is most efficient and can be expressed as:

$$P_j=P(\bigwedge_i \alpha_i-\alpha_j\geq 0\mid y,X)=P(\eta^{(j)}\geq 0\mid y,X),$$

where $\eta^{(j)}$ is the (N-1)x1 vector consisting of $\alpha_i-\alpha_j$ for $i=1,..,N$, $i\neq j$. Since the α_i's are Normally distributed, it follows that the marginal posterior of $\eta^{(j)}$ is the (N-1)-variate Normal distribution with means

$$E(\eta_i^{(j)} \mid y,X) = \hat{\alpha}_i - \hat{\alpha}_j,$$

and covariances

$$Cov(\eta_i^{(j)}, \eta_h^{(j)} \mid y,X) = \hat{\sigma}^2[(\bar{x}_i - \bar{x}_j)' S^{-1} (\bar{x}_h - \bar{x}_j) + \frac{1+\delta(i,h)}{T}],$$

for i,h = 1,..,N, $i \neq j$, $h \neq j$, where $\eta_i^{(j)}$ is $\alpha_i - \alpha_j$. Thus, P_j is the posterior probability mass located in the positive orthant of the (N-1)-dimensional space of $\eta^{(j)}$. It is possible to obtain analytical approximations to P_j, but we choose to perform Monte Carlo integration since the Monte Carlo draws used for calculating P_j can also be used for calculating $p(u_i^{rel} | y,X,u_j^{rel}=0)$. That is, for each j, we draw random vectors from the appropriate (N-1)-dimensional Normal distribution and count the proportion of draws which have all elements positive; this proportion is an estimate for P_j. Note that this procedure theoretically is very computationally intensive since it must be performed for all N firms. In practice, however, it will usually be the case that only a few firms have non-negligible P_j's. For this reason, we pursue the following strategy: the firms are ordered from the smallest $\hat{\alpha}_j$ to the largest. We start by computing P_1 (which is typically the largest), followed by P_2, etc. We stop computation when $\Sigma P_j > .999$. All subsequent P_j's are set to zero. Hence, the computational demands of this approach are much reduced.

$p(u_i^{rel} | y,X,u_j^{rel}=0)$ can be calculated as a by-product of the Monte Carlo integration procedure described in the previous paragraph. That is, $p(u_i^{rel} | y,X,u_j^{rel}=0)$ is equivalent to $p(\eta_i^{(j)} | y,X,\eta^{(j)} \geq 0)$, which is the appropriate marginal from the joint truncated Normal distribution. Hence, the accepted Monte Carlo draws used in calculating the P_j's automatically provide draws from $p(u_i^{rel} | y,X,u_j^{rel}=0)$. As log costs is the dependent variable, and interest centres on $\exp(-u_i^{rel})$, the Monte Carlo draws are transformed and used to plot the efficiency measure.

One more point should be noted before proceeding to the one-sided individual effects case. It is often of interest to allow for firm-specific inefficiency to depend upon some other variables. For example, for hospitals it is of interest to see if different organizational structures (eg. non-profit vs. for-profit) tend to imply different efficiency levels. However, variables such as for-profit status do not vary over time. If we were to introduce such time-invariant variables, then the kxk matrix S would be singular and the posterior would not be defined. Hence, for the standard individual effects case we do not allow for firm-specific efficiencies to depend on organizational structure or other firm characteristics.

Marginally Independent Efficiency Distribution (MIED) Model

The second Bayesian fixed effects model is still characterized by the absence of links between the individual effects, which, for a Bayesian, translates into marginal prior independence between the u_i's. In the preceding discussion of the standard model, a noninformative prior was used for α. However, we now use an informative prior. When dealing with stochastic frontiers, the individual effects, u_i, are measures of inefficiency and thus, by definition, are non-negative. This fact has motivated various classical maximum likelihood studies (eg. Pitt and Lee (1981) and Schmidt and Sickles (1984)) and will be at the basis of our models with proper distributions on the individual effects.

As before, equation (2) provides the basic model, and let $u=(u_1,..,u_N)'$ be the vector of firm specific efficiencies. The non-negativity of the u_i's can be thought of as prior information that the researcher should impose. We assume a particular one-sided prior distribution for u. We let u_i be independent of v_i and i.i.d. exponential with firm-specific mean λ_i. Note that the specification of a distribution for the u_i's allows us to talk of absolute inefficiencies, unlike the standard individual effects case. The parameters λ_i^{-1} are assumed to have independent exponential priors with means all equal to $-1/\ln(r^*)$ $(i=1,...N)$. The marginal prior of the efficiency of firm i, $r_i=\exp(-u_i)$, is given by
$p(r_i)= r_i^{-1}\, f_{IB}(-\ln(r_i)\,|\,1,\ 1,\ -\ln(r^*))$, where

$$f_{IB}(z\,|\,a,b,c) = \frac{\Gamma(a+b)}{c\Gamma(a)\Gamma(b)}\left(\frac{z}{c}\right)^{b-1}\left(1+\frac{z}{c}\right)^{-(a+b)},$$

denotes the density function of the three-parameter inverted Beta or Beta prime distribution (see Zellner (1971, p. 375-376)). Since r^* is the prior median efficiency, prior elicitation can be performed based on an easily understood quantity. As a result of the prior independence of the λ_i's, efficiencies are marginally prior independent. We stress that an explicit parameterization in terms of λ_i's is not formally required, as we could immediately start from the marginal Inverted Beta prior distributions on the u_i's. The corresponding Bayesian model then becomes:

$$p(y,u,\alpha_0,\beta,\sigma^{-2}\,|\,X)=p(y\,|\,X,u,\alpha_0,\beta,\sigma^{-2})p(u)p(\alpha_0,\beta,\sigma^{-2})$$
$$=c\sigma^2\prod_{i=1}^{N} f_N^T(y_i\,|\,X_i\beta+(\alpha_0+u_i)\iota_T,\sigma^2 I_N)f_{IB}(u_i\,|\,1,1,-\ln(r^*)) \tag{10}$$

However, introducing the incidental parameters λ_i and thus using a trivial hierarchical prior structure considerably facilitates the numerical analysis of this model through Gibbs sampling. In addition, it allows for a more direct comparison with the random effects models, to be introduced later.

In the absence of data corresponding to a particular firm f, the individual effect u_f will not be updated by the observations in the sample. Note that the latter result requires both prior independence between u_f and (u, α_0, β, σ^{-2}) as well as sampling independence over firms, which is assumed throughout. Therefore, in Bayesian fixed effects models the sample can not help us in predicting individual effects (efficiencies) of unobserved firms.

We wish to make inferences about α_0, β, σ^2 and u. It turns out that the joint posterior distribution is very difficult to work with. However, conditional posterior distributions have relatively simple forms. This suggests that a Gibbs sampler can be set up for this model.[1] In particular, conditional on u and λ^{-1} = (λ^{-1}... λ_N^{-1}), the posterior density of the frontier parameters and precision, σ^{-2}, has the usual Normal-Gamma form:

$$p(\alpha_0,\beta,\sigma^{-2}\,|\,y,X,u,\lambda^{-1})=p(\sigma^{-2}\,|\,y,X,u)p(\alpha_0,\beta\,|\,y,X,u,\sigma^{-2}), \tag{11}$$

where

$$p(\sigma^{-2}\,|\,y,X,u)=f_G(\sigma^{-2}\,|\,\frac{NT-k-1}{2},$$
$$\frac{1}{2}[y-(\iota_{NT}\vdots X)\binom{\alpha^*}{\beta^*}-(I_N\otimes\iota_T)u]'[y-(\iota_{NT}\vdots X)\binom{\alpha^*}{\beta^*}-(I_N\otimes\iota_T)u]), \tag{12}$$

and

$$p(\alpha_0,\beta\,|\,y,X,u,\sigma^{-2})=f_N^{k+1}\left(\begin{matrix}\alpha_0\\ \beta\end{matrix}\,\Big|\,\binom{\alpha^*}{\beta^*},\sigma^2\begin{bmatrix} NT & NT\bar{X}^{*\prime} \\ NT\bar{X}^* & X'X\end{bmatrix}^{-1}\right). \tag{13}$$

In the previous equations, $f_G(.|a,b)$ denotes the density function of the Gamma distribution with mean a/b and variance a/b^2, $y=(y_1...y_N)'$ an NTx1 vector, $X=(X_1'...X_N')'$ an NTxk matrix, and

---

[1]An introduction to the Gibbs sampler is given in Gelfand and Smith (1990) or Koop (1994). A discussion of the use of Gibbs sampling techniques in stochastic frontier models with cross-sectional data is given in Koop, Steel and Osiewalski (1994).

$$\begin{pmatrix}\alpha^{*}\\ \beta^{*}\end{pmatrix}=\begin{bmatrix}NT & NT\bar{X}^{*\prime}\\ NT\bar{X}^{*} & X'X\end{bmatrix}^{-1}\begin{bmatrix}NT(\bar{y}^{*}-\bar{u})\\ X'y-X'(I_N\otimes\iota_T)u\end{bmatrix}.$$

Here,

$$\bar{y}^{*}=\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}y_{it},$$
$$\bar{u}=\frac{1}{N}\sum_{i=1}^{N}u_i,$$

and,

$$\bar{X}^{*}=\frac{1}{NT}\sum_{i=1}^{N}\sum_{t=1}^{T}x_{it}.$$

The conditional posterior for the inefficiencies takes the form:

$$p(u\mid y,X,\alpha_0,\beta,\sigma^{-2},\lambda^{-1})\propto f_N^N(u\mid \bar{y}-(\iota_N\vdots\bar{X})\binom{\alpha_0}{\beta}-\frac{\sigma^2}{T}\lambda^{-1},\frac{\sigma^2}{T}I_N)\times I^{+}(u), \qquad (14)$$

where

$$\bar{y}=\begin{bmatrix}\bar{y}_1\\ \cdot\\ \cdot\\ \bar{y}_N\end{bmatrix},\quad \bar{X}=\begin{bmatrix}\bar{X}_1'\\ \cdot\\ \cdot\\ \bar{X}_N'\end{bmatrix},$$

and $I^{+}(u)$ is the indicator function for R_+^N. In other words, the u_i's are independently Normally distributed, but truncated to the positive orthant. The conditional posterior for λ^{-1} becomes:

$$p(\lambda^{-1}\mid y,X,u,\alpha_0,\beta,\sigma^{-2})=p(\lambda^{-1}\mid u)=\prod_{i=1}^{N}f_G(\lambda_i^{-1}\mid 2,u_i-\ln(r^{*})). \qquad (15)$$

A Gibbs sampler can be set up in terms of equations (11), (14) and (15). Random sampling from all these densities is standard. Note that, despite the high dimensionality of the problem ($2N+k+2=803$ in our problem), three steps suffice for each Gibbs draw. It is worth stressing that this specification assumes a separate efficiency distribution for each firm. Thus, conditionally upon the parameters describing the frontier, the u_i's are posterior independent and are only updated by the T observations for firm i, similarly to the α_i's in the SIE model.

2.2 Bayesian Random Effects Models

In this Subsection, we focus on models where the individual effects are in some way related across firms. Here we distinguish between two models depending on the way in which these links are implemented. In the usual panel context, one often considers a hierarchical prior for α where α is Normal with mean μ_α and variance σ_α^2. The parameters μ_α and σ_α^{-2} in turn have a Normal-Gamma prior distribution (see e.g. Box and Tiao (1973)). It can easily be seen that this model has the same type of structure as a classical Normal random effects model, which is commonly assumed in the panel data literature, as e.g. in Simar (1992). The resulting variance component model with an intra-class covariance structure is typically analyzed using generalized least squares in a classical framework (see e.g. Schmidt and Sickles (1984) and Simar (1992)), and can readily be treated using Gibbs sampling in a Bayesian context (see, eg., Gelfand and Smith (1990)). However, we are here dealing with individual effects, u_i, that are nonnegative by definition, and thus we will use hierarchical structures that reflect this property. For classical maximum likelihood estimation of random one-sided individual effects models, see Pitt and Lee (1981) and Schmidt and Sickles (1984).

Varying Efficiency Distribution (VED) Models

One might reasonably assume that the efficiency of hospitals with similar characteristics could be related. One way of implementing such links is to parameterize the mean inefficiencies λ_i, which were all independent in the MIED models, as $\exp(-w_i'\gamma)$, where w_i is an mx1 vector of exogenous variables, which do not have a time component, to be used in explaining firm-specific inefficiency (eg. non-profit status) and $\gamma=(\gamma_1...\gamma_m)'$ is an mx1 parameter vector, linking the firm effects across the sample. The parameter vector γ is assumed to have a proper prior, independent of the other parameters. All other assumptions are identical to the MIED model. The Bayesian model is now:

$$\begin{aligned} & p(y,u,\gamma,\alpha_0,\beta,\sigma^{-2} \mid X,W) \\ &= p(y \mid X,u,\gamma,\alpha_0,\beta,\sigma^{-2})\, p(u,\gamma \mid W)\, p(\alpha_0,\beta,\sigma^{-2}) \\ &= c\,\sigma^2 p(\gamma) \prod_{i=1}^{N} f_N^T(y_i \mid X_i\beta+(\alpha_0+u_i)\iota_T, \sigma^2 I_N)\, f_G(u_i \mid 1,\exp(w_i'\gamma)) \end{aligned} \tag{16}$$

where W = $(w_1....w_N)'$. We assume $w_{i1}=1$. With the Gibbs sampler we now have to deal with numerical integration in only N+k+m+2 dimensions (425 in our empirical application). However, this model poses an added difficulty as the full conditional of γ is not a standard distribution. In addition, the sampler now involves four steps. We wish to stress that our framework explicitly allows for functional links between X_i and w_i, given that we condition on X (and W) throughout. That is, we allow the individual effects to be correlated with the variables describing the frontier.

The conditional posterior for the frontier parameters is identical to that given previously in equations (11), (12) and (13). The conditional posterior for the inefficiencies takes the form:

$$p(u \mid y,X,\alpha_0,\beta,\sigma^{-2},\gamma) \propto f_N^N(u \mid \bar{y}-(\iota_N\!:\!\bar{X})\binom{\alpha_0}{\beta} - \frac{\sigma^2}{T}\zeta, \frac{\sigma^2}{T} I_N) \times I^+(u), \tag{17}$$

where $\zeta = (\exp(w_1'\gamma)...\exp(w_N'\gamma))'$. In other words, the u_i's are again independently Normally distributed, truncated to the positive orthant.

The conditional posterior for γ depends on the form of the prior for γ. In order to maintain comparability with the common efficiency distribution case, to be introduced later, we reparameterize as

$$\exp(w_i'\gamma) = \prod_{j=1}^{m} \phi_j^{w_{ij}},$$

where $\varphi_j = \exp(\gamma_j)$.[2] We express our prior information in terms of the φ_j's. In particular, we assume that they are, a priori, i.i.d. Gamma with parameters a_j

---

[2]We also assume all w_{ij}'s are nonnegative, which is true in the present application. If it were not true, we could add constants and subtract the appropriate term from γ_1.

and g_j, $j=1,..,m$. Despite the conditional prior independence of the u_i's, inefficiencies are marginally linked through the common parameter γ in the conditional mean.

The full conditionals for φ_1 and for $\varphi_2,..,\varphi_m$ do not depend on the data or on $(\alpha_0,\beta,\sigma^{-2})$, and can be written as:

$$p(\phi_1 \mid u,\phi_2,..,\phi_m)=f_G(\phi_1 \mid N+a_1,g_1+\sum_{i=1}^{N} u_i D_i), \tag{18}$$

where

$$D_i=\prod_{j=2}^{m} \phi_j^{w_{ij}}, \quad m>1,$$

and

$$p(\phi_2,..,\phi_m \mid u,\phi_1) \propto A(\phi_2,..,\phi_m)\exp(-\phi_1\sum_{i=1}^{N} u_i D_i), \tag{19}$$

where $A(\varphi_2,..,\varphi_m)$ is the product of m-1 independent Gamma densities for $\varphi_2,..,\varphi_m$. Each of these Gamma densities has parameters $a_j+N\ \bar{w}_j$ and g_j, where $\bar{w}_j = N^{-1}\sum_{i=1}^{N} w_{ij}, j=1,...m$.

The Gibbs sampler involves drawing from (11), (17), (18) and (19). Only (19) poses any difficulty. We use an independence Metropolis algorithm to draw from this conditional distribution (see Tierney (1991)). Like the Gibbs sampler, the Metropolis algorithm, is based on a Markov chain. A Markovian transition kernel drives the chain by generating candidate values for the next draw. These candidates are then either accepted with a certain probability, or rejected, in which case the chain remains at the current value. The independence Metropolis chain draws candidates independently and always from the same density, $\theta(.)$. So, on the l'th pass, this algorithm generates a candidate, φ^*, from $\theta(\varphi)$, where $\varphi=(\varphi_2...\varphi_m)'$. The random draw from (19), $\varphi^{(l)}$, is then either φ^* (with probability P_s), or the previous value, $\varphi^{(l-1)}$. Tierney stresses that this method works best if $\theta(\varphi)$ is a good approximation to the actual posterior. For a more detailed exposition of this method in the context of cost frontiers with a globally flexible functional form, see Koop, Osiewalski and Steel (1994).

Since φ is always non-negative, $\theta(.)$ should reflect this. Furthermore, it can be shown that all of the conditionals, based on (19), for φ_k (given the other φ_j, $j \neq k$), are unimodal.[3] A convenient one-sided unimodal distribution which allows for fat tails is the inverted Beta distribution, IB(e,b,c). We use this distribution independently for each of the elements of φ and take $c=1$. Note that this is a very flexible distribution with mode $\min\{0, (b-1)/(e+1)\}$ that allows for very fat tails through the choice of e. In practice, we work with m-1 independent $IB(e_j,b_j,1)$ distributions, $j=2,..,m$. The parameters e_j and b_j are calibrated on the basis of initial runs to ensure that the mode of $\theta(.)$ is near that of (19) and that the tails of $\theta(.)$ are at least as fat as those of (19). The probability of switching, P_s, is given by:

$$\ln(P_s) = \min(0, \phi_1 \sum_{i=1}^{N} [u_i(D_i^{(l-1)} - D_i^*)] + \sum_{j=2}^{m} [g_j(\phi_j^{(l-1)} - \phi_j^*) + (N\bar{w}_j + a_j - b_j)\ln(\frac{\phi_j^*}{\phi_j^{(l-1)}}) + (e_j + b_j)\ln(\frac{1+\phi_j^*}{1+\phi_j^{(l-1)}})]),$$

where D_i^* and $D_i^{(l-1)}$ are D_i evaluated at the candidate draw, φ^*, and the (l-1)'th value, $\varphi^{(l-1)}$, respectively. Finally, provided $N\bar{w}_j > b_j - a_j$ (which is satisfied in our application), the ratio of (19) and $\theta(\varphi)$ is bounded for any $\varphi \in R_+^{m-1}$, and thus we have uniform ergodicity of the Metropolis chain, implying the strongest form of convergence (Tierney (1991)).

Common Efficiency Distribution (CED) Model

In the previous VED model only the efficiencies of the firms with the same characteristics (as measured by w_i) are drawn from a common distribution. Here we will develop an important special case of the VED model, where $m=1$, and, since $w_{i1}=1$, this amounts to assuming that all individual effects are independent drawings from the same distribution. Thus, the links between firm effects will be even stronger than in the previous model.

We now assume that u_i is still independent of v_i and i.i.d. exponentially distributed with a common mean μ. Thus, the CED is a special case of the VED

---

[3]Formally, this is only true if $N\bar{w}_k > 1-a_k$. In our application, this unimodality result always holds.

model with m=1 and $\mu=\phi_1^{-1}$. The Bayesian model is given by (16) with $W=\iota_N$, and $\gamma=-\ln(\mu)$. The prior on μ^{-1} is now taken to be exponential with mean $-1/\ln(r^*)$, i.e. the same as the marginal prior for each λ_i^{-1} used to parameterize the MIED model. Therefore, the marginal distribution of r_i will be exactly the same as in the latter model, but the r_i's will no longer be prior independent. As μ now has the interpretation of a common mean efficiency, we will also be interested in inference on μ. For T=1 our CED model reduces exactly to the exponential model used for analyzing cross-sectional data in van den Broeck, Koop, Osiewalski and Steel (1994).

Again, the Gibbs sampler is a natural method to treat the numerical integration required for posterior and predictive inference. It will now be implemented by cyclical drawings from (11) and (14), where $\lambda^{-1}=\mu^{-1}\iota_N$, and from the full conditional for μ^{-1} which is:

$$p(\mu^{-1}\mid y,X,u,\alpha_0,\beta,\sigma^{-2})=p(\mu^{-1}\mid u)=f_G(\mu^{-1}\mid (N+1),N\bar{u}-\ln(r^*)). \qquad (20)$$

We remind the reader that the CED model is a special case of the VED model, and we would like to specify a prior on the latter that is consistent with the prior assumed for the CED model. In the VED model each ϕ_j had a Gamma prior distribution with parameters a_j and g_j, j=1,...m. To ensure prior consistency with the CED model, we set $g_1=-\ln(r^*)$, and $g_j=1$ for j=2,..,m and $a_j=1$ for j=1,..,m. In other words, we centre the prior over the CED model with the same hyperparameter, r^*, which is no longer the prior median if m>1. In the VED case, $P(r_i<=r^*) = E[(1+\prod_{j=2}^{m}\phi_j)^{-1}]$, where the expectation is with respect to the prior of $\phi_2,...,\phi_m$. Applying Jensen's inequality, we conclude that $P(r_i<=r^*) > 0.5$, i.e. the prior median efficiency is less than r^* whenever m>1.

Tables 2 and 3 and Figure 1 illustrate some properties of the prior for the varying and common efficiency distribution cases. The latter is a special case of the former with m=1. Table 2 contains prior efficiency means and standard deviations for various values of r^* for $a_j=g_j=1$ (j=2,..,m). Although changing r^* changes the prior moments, a comparison of m=1 with m=4 for $r^*=0.8$ in Figure 1 indicates that the priors have roughly similar properties.

Table 3 investigates the effect of changing a_j and g_j (j=2,..,m) for the m=4 case. We keep $a_j=g_j$, but let them take on a common value different from 1.

This implies that the φ_j's still have mean 1 (and hence are centred over the m=1 model), but allows their prior variances to take on values different from 1. It can be seen that changing these prior variances has relatively little effect on the prior efficiency. Hence, in our empirical work, we set $a_j=g_j=1$ (j=2,..,m), which is judged a reasonable value, and do not investigate other values for these prior hyperparameters. Clearly, as a_j and g_j grow for j=2,..,m, the prior for φ_j (j>1) will become tighter around one and prior efficiency for the case m=4 will tend to that with m=1 (which can be derived from the model with m>1 by restricting all φ_j (j>1) to be one). Marginally, the prior moments for the efficiency in the MIED model will be the same as in the CED (m=1) model, but for the former the conditional prior independence is preserved in the marginal prior for the r_i's.

Figure 1 plots the marginal prior density of efficiency for the m=1 ($r^*=.8$) and m=4 ($r^*=.8$, $a_j=g_j=1$ for j=2,..,m) cases. In our prior sensitivity analysis, to be discussed in Section 4, we find that posterior results for m=1 are extremely robust for the random effect models, even to enormous changes in r^*. In view of this, the small differences in prior between the m=1 and m=4 cases will undoubtedly have negligible empirical consequences. In contrast, the choice of r^* is found to be much more important in the MIED model, in line with the fixed effects structure of that model. In addition, Figure 1 contains the continuous part of the implied prior on the relative efficiency r_i^{rel} for the SIE model (with arbitrary scaling), from which it is obvious that this model corresponds to a very strong prior belief in low efficiency. The proper priors, on the other hand, convey a genuine sense of lack of strong prior information. The particular U-shaped form of their densities reflects the thick tail of the marginal prior on u_i, which is evident from (10) in the case of the MIED model.

2.3 Common Efficiency (CE) Model

An extreme case of linking the individual efficiencies would be to assume that they are all exactly the same, i.e. $u_i=z$ (i=1,...N), where the prior on z can be degenerate at a particular value (e.g. zero to reflect full efficiency) or can have any other form. Since all u_i's are equal in this model, we can no longer talk of individual effects. From an economic point of view, such a restrictive model with $m=1$ and $\mu=\phi_1^{-1}$. The Bayesian model is given by (16) with $W=\iota_N$, and $\gamma=-\ln(\mu)$. The prior on μ^{-1} is now taken to be exponential with mean $-1/\ln(r^*)$, i.e. the same as the marginal prior for each λ_i^{-1} used to parameterize the MIED model. Therefore, the marginal distribution of r_i will be exactly the same as in the latter model, but the r_i's will no longer be prior independent. As μ now has the interpretation of a common mean efficiency, we will also be interested in inference on μ. For $T=1$ our CED model reduces exactly to the exponential model used for analyzing cross-sectional data in van den Broeck, Koop, Osiewalski and Steel (1994).

Again, the Gibbs sampler is a natural method to treat the numerical integration required for posterior and predictive inference. It will now be implemented by cyclical drawings from (11) and (14), where $\lambda^{-1}=\mu^{-1}\iota_N$, and from the full conditional for μ^{-1} which is:

$$p(\mu^{-1}\mid y,X,u,\alpha_0,\beta,\sigma^{-2})=p(\mu^{-1}\mid u)=f_G(\mu^{-1}\mid (N+1),N\bar{u}-\ln(r^*)). \qquad (20)$$

We remind the reader that the CED model is a special case of the VED model, and we would like to specify a prior on the latter that is consistent with the prior assumed for the CED model. In the VED model each ϕ_j had a Gamma prior distribution with parameters a_j and g_j, $j=1,\ldots m$. To ensure prior consistency with the CED model, we set $g_1=-\ln(r^*)$, and $g_j=1$ for $j=2,..,m$ and $a_j=1$ for $j=1,..,m$. In other words, we centre the prior over the CED model with the same hyperparameter, r^*, which is no longer the prior median if $m>1$. In the VED case, $P(r_i<=r^*) = E[(1+\prod_{j=2}^{m}\phi_j)^{-1}]$, where the expectation is with respect to the prior of $\phi_2,\ldots,\phi_m$. Applying Jensen's inequality, we conclude that $P(r_i<=r^*) > 0.5$, i.e. the prior median efficiency is less than r^* whenever $m>1$.

Tables 2 and 3 and Figure 1 illustrate some properties of the prior for the varying and common efficiency distribution cases. The latter is a special case of the former with $m=1$. Table 2 contains prior efficiency means and standard deviations for various values of r^* for $a_j=g_j=1$ $(j=2,..,m)$. Although changing r^* changes the prior moments, a comparison of $m=1$ with $m=4$ for $r^*=0.8$ in Figure 1 indicates that the priors have roughly similar properties.

Table 3 investigates the effect of changing a_j and g_j $(j=2,..,m)$ for the $m=4$ case. We keep $a_j=g_j$, but let them take on a common value different from 1.

This implies that the φ_j's still have mean 1 (and hence are centred over the m=1 model), but allows their prior variances to take on values different from 1. It can be seen that changing these prior variances has relatively little effect on the prior efficiency. Hence, in our empirical work, we set $a_j=g_j=1$ $(j=2,..,m)$, which is judged a reasonable value, and do not investigate other values for these prior hyperparameters. Clearly, as a_j and g_j grow for $j=2,..,m$, the prior for φ_j $(j>1)$ will become tighter around one and prior efficiency for the case $m=4$ will tend to that with $m=1$ (which can be derived from the model with $m>1$ by restricting all φ_j $(j>1)$ to be one). Marginally, the prior moments for the efficiency in the MIED model will be the same as in the CED ($m=1$) model, but for the former the conditional prior independence is preserved in the marginal prior for the r_i's.

Figure 1 plots the marginal prior density of efficiency for the $m=1$ $(r^*=.8)$ and $m=4$ $(r^*=.8,\ a_j=g_j=1$ for $j=2,..,m)$ cases. In our prior sensitivity analysis, to be discussed in Section 4, we find that posterior results for $m=1$ are extremely robust for the random effect models, even to enormous changes in r^*. In view of this, the small differences in prior between the $m=1$ and $m=4$ cases will undoubtedly have negligible empirical consequences. In contrast, the choice of r^* is found to be much more important in the MIED model, in line with the fixed effects structure of that model. In addition, Figure 1 contains the continuous part of the implied prior on the relative efficiency r_i^{rel} for the SIE model (with arbitrary scaling), from which it is obvious that this model corresponds to a very strong prior belief in low efficiency. The proper priors, on the other hand, convey a genuine sense of lack of strong prior information. The particular U-shaped form of their densities reflects the thick tail of the marginal prior on u_i, which is evident from (10) in the case of the MIED model.

2.3 Common Efficiency (CE) Model

An extreme case of linking the individual efficiencies would be to assume that they are all exactly the same, i.e. $u_i=z$ $(i=1,...N)$, where the prior on z can be degenerate at a particular value (e.g. zero to reflect full efficiency) or can have any other form. Since all u_i's are equal in this model, we can no longer talk of individual effects. From an economic point of view, such a restrictive assumption would typically be far too dogmatic in most applications. Statistically, the common effect z does not differ across individuals and is thus indistinguishable from the intercept α_0. Note that the sampling model is now sufficiently parameterized in terms of $\kappa=\alpha_0+z$, β and σ^{-2}. Assuming a prior of the form $p(z, \alpha_0, \beta, \sigma^{-2})=p(z)\, p(\alpha_0, \beta, \sigma^{-2})$ with the same improper prior on the parameters of the frontier as used before, the product structure is preserved under the parameter transformation from α_0 to κ. Thus, our Bayesian model becomes:

$$\begin{aligned} p(y,z,\kappa,\beta,\sigma^{-2}\mid X) &= p(y\mid X,\kappa,\beta,\sigma^{-2})p(z,\kappa,\beta,\sigma^{-2}) \\ &= c\,\sigma^{2}p(z)\prod_{i=1}^{N} f_N^T(y_i\mid X_i\beta+\kappa\iota_T,\sigma^{2}I_N) \end{aligned} \tag{21}$$

For a Bayesian, the lack of identification translates into an absence of updating of the marginal prior on z, as is clearly shown in (21). As a benchmark case, this model, which does not involve any numerical integration, will also be examined.

3. Hospital Cost Function Estimation

The theory of the multiple-product firm implies that a firm's costs should depend upon the quantity of each output produced as well as the input prices faced by the firm. Given data on outputs and input prices, the researcher can select a functional form for the cost frontier and estimate its parameters. Hospital cost function estimation poses problems which render it difficult to apply such a strategy in a straightforward manner.[4] A modern hospital produces a myriad of outputs that are hard to quantify.[5] For most standard

---

[4]Cowing, Holtmann and Powers (1983) provides a discussion of some of the difficulties inherent in hospital cost function estimation.

[5]Most theories of the firm imply that cost minimization should be a reasonable objective function even for non-profit or government-run institutions. Empirical evidence for or against the assumption of cost minimization is scanty. One exception is Eakin and Kniesner (1988), which estimates a long-run cost function (using 1975-76 data) that allows for systematic allocative inefficiency and rejects the assumption of cost minimization. However, since estimated differences between shadow and observed marginal costs are small, the authors conclude that use of traditional minimum cost functions may yield fairly accurate estimates of output concepts, such as economies of scale and scope.

functional forms (eg. the translog), the large number of outputs causes the number of parameters to be estimated to be very large. As a result, researchers have worked with highly aggregated data. Early work in this area avoided such flexible functional forms and worked with ad hoc (usually linear) specifications, where the class of explanatory variables was expanded beyond that implied by economic theory (see Breyer (1987) for a discussion). Much of the recent work (see Vita (1990) and Granneman, Brown and Pauly (1986)) has criticized these ad hoc specifications and worked with flexible functional forms such as the translog. In this paper, we intend to follow the path of these latter authors and work with highly aggregated data and a translog functional form. In addition, we adopt the stochastic frontier framework described in the previous Section.

Breyer (1987), in a survey of the hospital cost function literature, argues that the true "output" of a hospital is improvement in patient health. Defined in this way, "output" is impossible to measure, so Breyer recommends using observable intermediate products as proxies for output. In particular, he suggests three important hospital output dimensions that can easily be measured: i) Number of cases (as a proxy for medical services); ii) Number of inpatient days (as a proxy for nursing, accommodation and other "hotel" services); and iii) Number of beds (to satisfy an option demand for hospital services).[6] In the present paper, we use these three variables as measures of output. In addition, we include the number of outpatient visits and a case mix index as other aspects of outputs that are included in our cost frontier.

In order to estimate a cost frontier, it is important that all firms face the same technology. If this is the case, then deviations from the cost frontier can be interpreted as evidence of inefficiency. However, in the hospital cost function literature, it is often argued that different types of hospitals have different technologies. For this reasons, it is important to have a panel of hospitals that is as homogeneous as possible. Hence, we omit from our panel all teaching hospitals, since such hospitals are typically thought to behave differently from non-teaching hospitals. Furthermore, it is possible that apparent inefficiencies are due to certain hospitals providing a higher quality of service that is not captured by our output measures. For instance, hospitals which have unusually

---

[6]It has been argued by some researchers that the number of beds is better considered as a particular type of capital, this should be kept in mind when considering our empirical results.

high ratios of employees to patients may be inefficient, or they may be offering a higher quality of service -- we cannot know which. In light of this, we have deleted from our panel any hospital which had an unusually high or low employee to patient ratio in any year.[7] In this manner, we have ensured that the panel of hospitals under consideration is relatively homogeneous.

In our cost frontier, we also include a measure of capital stock (total fixed assets). The fact that such a variable is included means that all our results should be interpreted as applying in the short run.

In terms of inputs, labour is predominant. Our data source contains only one aggregate wage index. The other significant input in the hospital technology is general materials and supplies. Unfortunately, we have no measurements on this. However, given the wholesale buying power of most hospitals, it is reasonable to assume that the price of materials is fairly constant across hospitals. Hence, we treat the price of materials as a constant. Such a treatment is undoubtedly reasonable cross-sectionally, but is not reasonable over time. Thus, we add a time trend and a time trend squared as explanatory variables in our cost frontier to try to capture the missing time dimension of the price of materials or other dynamics that are not modelled explicitly.

With this data, we can specify a standard cost function where costs depend on 5 output categories, one capital stock, two input prices, a time trend and a time trend squared. We choose a translog specification, and impose linear homogeneity in prices. The imposition of linear homogeneity allows us to normalize with respect to the price of materials (which is a constant), yielding the resulting cost frontier:

$$\begin{aligned} \ln C = \alpha_0 + \sum_{i=1}^{5}\beta_i \ln Y_i + \beta_6 \ln P + \sum_{i=1}^{5}\sum_{j=i}^{5}\psi_{ij}\ln Y_i \ln Y_j + \\ \beta_7(\ln P)^2 + \sum_{i=1}^{5}\beta_{7+i}\ln Y_i \ln P + \beta_{13}\ln K + \sum_{i=1}^{5}\beta_{13+i}\ln Y_i \ln K \\ + \beta_{19}\ln P \ln K + \beta_{20}(\ln K)^2 + \beta_{21}t + \beta_{22}t^2, \end{aligned} \tag{22}$$

where ψ_{ij} are the remaining 15 elements of β. A brief description of the

[7]More precisely, we have two variables: one of which measures the ratio of clinical workers to average daily census, the other which measures the ratio of nonclinical workers to average daily census. If either of these variables is more than one standard deviation from the mean for any year, then the hospital is deleted from our panel.

variables is given in Table 1, and more detail is provided in the Data Appendix. The explanatory variables (w_{ij}) for the varying efficiency distribution model are taken to be the ownership status (i.e. non-profit, for-profit, or government) and the Herfindahl index averaged over time.

4. Empirical Results

In this section we discuss our empirical findings for the models considered in Section 2: the standard individual effects (SIE) model, the marginally independent efficiency distribution (MIED) model, the varying efficiency distribution (VED) model, the common efficiency distribution (CED) model, and, finally, the common efficiency (CE) model. We remind the reader that the first two of these models are Bayesian fixed effects models, the second pair constitute Bayesian random effects models, and the fifth model does not allow for individual effects. These models are arranged in order of increasing prior links between the firm-specific efficiencies. The findings from these models are discussed and compared in different subsections containing: i) properties of the frontier, ii) properties of the efficiencies, iii) a sensitivity analysis with respect to the chosen prior, and iv) explaining hospital-specific efficiencies. It is worth emphasizing that prior elicitation is done on the basis of r^*, which is the prior median efficiency in the MIED and CED models, and we set this prior hyperparameter to 0.8 throughout our analysis, except for the discussion of prior sensitivity.

Of these models, the CE model is the simplest in that it does not require numerical integration, the SIE will be treated using direct Monte Carlo drawings, and the numerical integration in the remaining three models can be done efficiently through Gibbs sampling, as described in Section 2. Building upon our findings in Koop, Steel and Osiewalski (1994) and Koop, Osiewalski and Steel (1994), we have used a parallel implementation of Gibbs sampling, where we conduct M runs, each containing L passes, and keep only the Lth pass of each run. This strategy can be very effective in breaking serial correlation, induced by the Markovian structure of Gibbs sampling, and prevents the sampler from being stuck in a region containing small posterior probability. We have used $M=1000$ throughout, whereas $L=50$ for all the results except for the MIED model with very low values of r^*, where larger L was required for full convergence.

Properties of the frontier

Since our focus is on hospital efficiency analysis, we will only briefly discuss the properties of the frontier. Instead of presenting posterior means and standard deviations for $k=38$ frontier parameters, Tables 4, 5 and 6 present posterior means and standard deviations of cost elasticities with respect to P, Y_i ($i=1,..,5$), and K, for the minimum, median, and maximum cost firms in our sample, respectively. Note that, with a few exceptions, all of these elasticities are consistent with the regularity conditions implied by economic theory. The exceptions typically occur for extreme firms and for the median firm regularity conditions are only violated with respect to Y_3, which could also be considered a type of capital stock. Given the weaknesses of our input price data and the fact that our output measures are not really true output measures, but rather reflect aspects of the true unobservable outputs, we feel that the properties of our frontier are quite satisfactory. Note that economic theory, apart from requiring the cost function to be increasing in Y_i ($i=1,..,5$) and P, also imposes concavity and homogeneity of degree one in P. The latter is assured through the normalization of (22), whereas the former translates into $\beta_7<0$. Posterior means (standard deviations) of β_7 for the SIE, MIED, VED, CED and CE models are -1.08 (0.23), -0.80 (0.21), -0.81 (0.22), -0.79 (0.21) and -0.59 (0.21) respectively, indicating that concavity is satisfies with virtual certainty.

The one-sided efficiency distribution models, MIED, VED and CED, exhibit fewer violations of regularity conditions than the SIE model. In general, the one-sided models yield very similar frontiers. The frontier for the SIE model leads to elasticities which tend to be much smaller than those of the one-sided efficiency models. As we shall show shortly, the inefficiencies of the SIE model tend to be larger and differ much more across firms than with the other individual effects models. Loosely speaking, the type of prior information on the inefficiency term in the standard model causes more of the variation in costs to be picked up as inefficiency. From the elasticities corresponding to the benchmark CE model, we conclude that eliminating individual effects (inefficiencies) seriously affects the inference on the frontier. Therefore, we would not recommend use of the CE model, even in those cases where one is only interested in properties of the cost frontier.

Properties of the Efficiency Measures

The Bayesian techniques used in this paper yield exact posterior distributions for the efficiency measures of each of our N hospitals. These are probably of greatest interest for policy purposes. However, space precludes a detailed presentation of efficiencies for each hospital. Instead, we present a detailed analysis of three firms which have minimum, median and maximum efficiency in a certain sense, and for the random effects models we consider an "average" hospital. The notion of the efficiency of an "average" or "typical" out-of-sample firm, r_f, is discussed in van den Broeck, Koop, Osiewalski and Steel (1994). Essentially, r_f is the efficiency of a hypothetical unobserved hospital. The posterior distribution of r_f is given by:

$$p(r_f \mid y,X) = r_f^{-1} \int_0^{\infty} f_G(-\ln(r_f) \mid 1, \mu^{-1}) p(\mu^{-1} \mid y,X) d\mu^{-1},$$

and by

$$p(r_f \mid y,X) = r_f^{-1} \int f_G(-\ln(r_f) \mid 1, \prod_{j=1}^{m} \phi_j^{\bar{w}_j}) p(\phi_1, .. \phi_m \mid y,X) d\phi_1 .. d\phi_m,$$

for the CED and VED models, respectively. For the latter model, we have characterized the "typical" firm as having average values for the w_{ij}'s, i.e. $\bar{w}_j$. Note that this implies a somewhat artificial "average" hospital since some of the w_{ij}'s are dummy variables. In other words, $p(r_f|y,X)$ is posterior to the data on all observed firms, but prior to the unobserved output for hypothetical firm f. The updating of r_f is only possible due to prior links between efficiencies, and hence considering r_f does not make sense for the fixed effects models.

We select our minimum/median/maximum efficient firms as having maximum/median/minimum values for $\hat{\alpha}_i$ in the standard model (see equation (6)). It is worth stressing that the efficiency of a particular hospital is a random variable. Hence, we cannot unambiguously say one hospital is more efficient than another. So, these minimum/median/maximum efficiencies ($r_{min}, r_{med}, r_{max}$) should only be considered as corresponding to firms which are probably very inefficient/about average/very efficient, respectively.

Figure 2 plots the marginal posterior density of r_f for the CED and VED cases. It can be seen that the plots are almost identical. If we compare Figure 2 with the prior for the one-sided efficiency distribution models in Figure 1, it can clearly be seen that the data rule out very low efficiencies. The posterior means (standard deviations) of r_f are 0.852 (0.128) for VED and 0.856 (0.125) for the CED model. As inference on r_f in the random effects models is conducted without knowledge of y_f and X_f, its distribution is much more spread out than the posterior distributions of firm-specific efficiencies. The distribution of r_f depicted in Figure 2 also captures the uncertainty about the type of hospital. For example, firm f could resemble the least efficient firm, but could also be like the most efficient one.

Figures 3 through 6 plot the posterior distributions of r_{min}^{rel}, r_{med}^{rel}, and r_{max}^{rel} for the SIE, and of r_{min}, r_{med} and r_{max} for the one-sided efficiency distribution models, respectively. Since Figures 4, 5 and 6 are almost identical, we will focus on comparing CED with SIE (Figure 3). Remember that, for the SIE model, efficiency is measured relative to the most efficient firm. It turns out that one firm is almost certainly the most efficient firm (i.e. the posterior probability that it is most efficient is .9997). For this reason the distribution of r_{max}^{rel} is a point mass at one. However, even for the CED model (which allows for the calculation of absolute efficiency), r_{max} has all of its probability mass very close to one. It is with regard to r_{med} and r_{min} that major differences occur. In particular, posterior means are much smaller for the standard model than for the other models (0.23 vs. around 0.70 for r_{min}^{rel} and r_{min}, respectively). This behaviour is consistent with the important difference between the SIE and one-sided distribution models. The implied prior on r_i^{rel} (see Figure 1) simply does not rule out very low relative efficiencies for many firms. From the fact that the posterior results for the MIED model (with $r^*=0.8$) are quite close to those of the random effects models, we infer that it is not the fixed effects nature that induces the SIE model to behave so differently from the rest.

If we replace the one-sided prior distribution on the inefficiencies with $r^*=0.8$ by the improper prior structure of the SIE model, we tend to substantially decrease the hospital efficiencies. Overall, the average posterior mean efficiency is 0.47 for the SIE model and around 0.85 for the one-sided efficiency distribution models. However, the differences go beyond merely decreasing the efficiency of each hospital; in many cases the ranking of efficiencies changes. The Spearman rank correlation between the N-vector of posterior means of hospital-specific efficiencies for the SIE and CED cases is only 0.43.

To illustrate the fact that the standard individual effects model considers relative, rather than absolute, measures of efficiency, we have eliminated the firm with the highest value of $\hat{\alpha}_i$ (which was most efficient with probability 0.9997). Then we need to consider 6 candidates for most efficient firm in order to capture a probability mass of at least 0.999, and the average posterior mean efficiency jumps to 0.55. Using the CED model on this reduced sample leads to an average posterior mean efficiency of 0.85, which hardly differs from that with the full sample.

The results from the one-sided models are, in our subjective opinion, much more reasonable than those of the SIE model, both in terms of the efficiency measures (is it reasonable that, as the SIE model would have, there are many hospitals with efficiencies less than 30% of the most efficient one?) and the frontier itself (fewer suggestions that regularity conditions are violated for the one-sided cases). Thus, we would advise against use of the SIE model. Therefore, in the rest of the discussion of our empirical results, we will primarily concentrate on the one-sided efficiency distribution models.

Even for the one-sided distribution models, the inference on firm-specific efficiencies suggests that cost-minimization is not achieved by many of the hospitals in our sample. Our analysis can not tell us, however, whether this fact is due to managerial error or to a different nature of the objective function.

Prior Sensitivity Analysis

The one-sided individual effects models are based upon proper priors for efficiencies. We believe that we have elicited quite reasonable priors, but it is always important to carry out a sensitivity analysis to see if the choice of prior hyperparameters (in our case r^*) has an important effect on our results. This prior hyperparameter r^* has the interpretation of the prior median efficiency for both the MIED and CED models. Let us compare these two models, as this enables us to isolate the effect of prior independence. For our previous discussions, we have set $r^*=.8$, implying that we assign a prior probability of

0.5 to any hospital being less than 80 per cent efficient.

Figure 7 displays the average of posterior mean efficiencies over the 382 hospitals within the sample, say $\bar{r}$, as a function of r^*. The striking feature of these graphs is that $\bar{r}$ (r^*) is virtually constant over the whole range from 0.01 until 0.99 for the CED model. It seems the random effect nature of this model links the efficiencies sufficiently to ensure that the data dominate the prior information. In sharp contrast, the fixed effects MIED model does not lead to such robustness, as $\bar{r}$ varies substantially with r^*. The independence assumption inherent in this model, combined with the small value of T, implies that the data information on each individual efficiency is much weaker than in the CED case. If we take a value of r^* in line with the data, say $r^*=0.8$, then both models lead to virtually the same inference on firm efficiencies, but if we deviate from such a prior median efficiency, differences between the models grow. As r^* becomes very small, the MIED model will have marginal priors on the efficiencies that are still proper, but will start to look like the marginal priors on relative efficiencies for the SIE model. As both models are fixed effects models and their only difference lies in the form of the marginal priors, we find, indeed, that results for the MIED model with $r^*=0.01$ are relatively close to the SIE model.

To further illustrate the robustness of our results for the CED model to extreme changes in our prior, Figure 8 plots the posterior distribution of r_f for $r^*=.01$, .80 and .99. It can be seen that the plots are virtually indistinguishable. Posterior means of r_f for these three values of r^* are 0.843 ,0.856, and 0.857 and posterior standard deviations 0.133, 0.125, and 0.124. This is strong evidence that results reported for the CED case are not dependent on the specific values of the prior hyperparameter.

Explaining Efficiencies

The VED model explicitly allows for hospital characteristics to affect efficiencies. We have considered a number of candidate variables for W. The characteristics we finally retain are non-profit, for-profit or government-run status and a Herfindahl index, which is averaged over time since the w_{ij} variables parameterize efficiencies which are assumed to be constant over time.

However, we can also use the other individual effects models to shed some light on factors which may affect efficiency. Table 7 presents average posterior means and standard deviations of individual efficiencies for for-profit, non-profit and government-run hospitals. Once again, the one-sided efficiency cases yield almost identical results for $r^{*}=0.8$. However, results for the SIE model are different. All models find that government-run hospitals are more efficient, but they disagree over whether non-profit hospitals tend to be more efficient than for-profit hospitals. Relative to the average posterior standard deviation of the efficiencies, the differences between types of hospitals is small. Overall, we cannot conclude that there is a pronounced systematic relationship between hospital organization and efficiency.

The varying efficiency distribution model allows us to investigate directly the effect of the w_i's on efficiency. Posterior means and standard deviations for the elements of γ are given in Table 8, in the order: intercept, non-profit, for-profit and average Herfindahl index. To aid in interpretation, note that the variables are normalized so that each element of w_i has mean 1 and that 70% of the hospitals in the panel are non-profit, 17% for-profit and 13% government-run. Thus, the value of w_{i2} for a non-profit hospital is 1.43 and w_{i3} for a for-profit hospital is 5.88. The posterior means of γ given in Table 7 are consistent with government-run hospitals being most efficient followed by non-profit and for-profit hospitals. The posterior mean of γ_4, the coefficient on the Herfindahl index, indicates that a less competitive environment tends to decrease efficiency. However, all these findings must be qualified since posterior standard deviations are large compared to means. Aside from γ_1 (which corresponds to the intercept), all of the posterior means are close to zero (relative to their standard deviation). An approximate Bayesian Highest Posterior Density interval test (see Zellner (1971), p. 298-302) strongly favours the CED model. Remember that the CED model is equal to the VED model with $\gamma_2=..=\gamma_m=0$. If the posterior for the γ_j's (j=2,..,m) is approximately Normal, then the distribution of the inner product of the standardized γ_j's is approximately $\chi^2(m-1)$. For our data, this quantity, evaluated at zero, is 1.94. This clearly indicates that the data do not provide evidence against the hypothesis that the γ_j's are 0 (j=2,..,m).

In Koop, Osiewalski and Steel (1994), a predictive measure of lack of fit is developed, a detailed justification of which is given therein. If we let $\epsilon_{ft}=u_f+v_{ft}$, then we advocate using $E(\epsilon_{ft}{}^2|y,X) = E(\sigma^2 + 2\mu^2|y,X)$ as a measure of fit for the CED model. For the VED model, $\exp(-w_i'\gamma)$ is analogous to μ, so we use $E(\sigma^2 + 2\exp(-2\iota_m'\gamma)|y,X,W)$, where ι_m is the average of the w_i's, which we use for w_f corresponding to the average firm. The measure of fit is 0.065 for the VED and 0.061 for the CED model.[8] In other words, in terms of our measure of fit, the VED model performs worse, as a result of the uncertainty in estimating γ. We cannot construct a comparable measure for the other models. However, note the the posterior mean of σ^2 is 0.0034 for the SIE model, 0.0042 for the MIED and 0.0043 for both random effects models. The fact that the prior of the SIE model strongly favours low efficiencies implies that measurement error is smaller for this latter model, as much more of the total error is allocated to inefficiency. At the other extreme, in the CE model all of the stochastics is, by definition, attributed to the measurement error term, which results in the posterior mean of σ^2 being 0.0145.

Figure 9 plots the posterior density of r_f for non-profit, for-profit and government-run hospitals based on the VED model. Corresponding means (standard deviations) are 0.859 (0.125), 0.764 (0.208) and 0.875 (0.135) for non-profit, for-profit and government-run hospitals, respectively. This merely reinforces our previous conclusions, *viz.* that we cannot with any confidence conclude that the type of hospital organization can help improve efficiency.

5. Conclusion

In this paper we have described and analyzed Bayesian models for inference on firm-specific efficiencies. We show how, by using different prior structures, we can derive Bayesian analogues to the classical fixed and random individual effects models. i) The fixed effects models are characterized by the absence of links between individual effects, and thus do not require a hierarchical prior structure. Within this class, we define the standard individual effects (SIE) model, which puts an improper uniform prior on the firm-specific intercepts and measures relative efficiencies as differences between the intercepts, and the marginally independent efficiency distribution (MIED) model, where independent

[8]Note that it is not necessary for this measure of fit to be smaller for the more flexible varying efficiency distribution model.

proper one-sided priors are used for individual effects, and efficiencies are thus defined in absolute terms. ii) Bayesian random effects models do link individual effects through the hierarchical structure of the prior, parameterizing the N effects in terms of a small number $m << N$ of additional parameters. The common efficiency distribution (CED) model takes $m=1$ and assigns a common exponential prior distribution to efficiencies; the varying efficiency distribution (VED) model allows the mean of the exponential prior to vary according to m-1 hospital characteristics.

The seemingly innocuous flat prior on individual intercepts associated with the SIE model enables us to reinterpret many classical results in a Bayesian context (see e.g. (4)-(7)). However, adoption of this model necessarily implies a strong prior belief in low firm efficiencies, which is very different from those commonly held. Furthermore, the fixed effects nature of this model makes it difficult for the sample to correct this prior information when T is small. For this reason, we would advocate using the one-sided individual effects models for inference on efficiencies in a stochastic frontier context. Whereas the MIED model allows us to capture our prior beliefs through any proper distribution that we judge reasonable, its fixed effects structure still implies a large sensitivity to the choice of the particular prior when T is not large. Furthermore, the lack of links between individual effects inherent in fixed effects models make prediction of these effects for unobserved firms a useless exercise.

We apply our methods to a panel of U.S. hospitals and obtain reasonable results for both random effects models, which display an impressive robustness with respect to large changes in our prior hyperparameter. One finding that is common to all the individual effects models examined is that variables reflecting hospital organization (i.e. non-profit, for-profit or government-run status) do not help explain hospital efficiencies. Even when explicitly modeled in the VED case, we have to conclude that ownership status as well as many other hospital characteristics (e.g. Herfindahl index) fail to give a deeper understanding of the substantial variability of efficiencies found in the sample.

References

Aigner, D., Lovell, C. and Schmidt, P. (1977). "Formulation and estimation of stochastic frontier production function models," **Journal of Econometrics, 6, 21-37.**

Box, G.E.P. and Tiao, G.C. (1973). **Bayesian Inference in Statistical Analysis,** Addison-Wesley: Reading.

Breyer, F. (1987). "The specification of a hospital cost function: A comment on the recent literature," **Journal of Health Economics, 6, 147-157.**

van den Broeck, J., Koop, G., Osiewalski, J. and Steel, M.F.J. (1994). "Stochastic frontier models: A Bayesian perspective," **Journal of Econometrics, 61,** 273-303.

Caves, D., Christensen, L. and Tretheway, M. (1980). "Flexible cost functions for multiproduct firms," **Review of Economics and Statistics,** 62, 477-481.

Cowing, T., Holtmann, A. and Powers, S. (1983). "Hospital cost analysis: A survey and evaluation of recent studies," **Advances in Health Economics and Health Services Research,** 4, 257-303.

Eakin, B. and Kniesner, T. (1988). " Estimating a non-minimum cost function for hospitals," **Southern Economic Journal,** 54, 583-597.

Engle, R.F., Hendry, D.F. and Richard, J.F. (1983). "Exogeneity," **Econometrica,** 51, 277-304.

Farrell, M. (1957). "The measurement of productive efficiency," **Journal of the Royal Statistical Society, A,** 120, 253-281.

Garnick, D., Luft, H., Robinson, J. and Tetreault, J. (1987). "Appropriate measures of hospital market areas," **Health Services Research,** 22, 69-89.

Gelfand, A. and Smith, A.F.M. (1990). "Sampling-based approaches to calculating marginal densities," **Journal of the American Statistical Association, 85,** 398-409.

Grannemann, T., Brown, R. and Pauly, M. (1986). "Estimating hospital costs: A-multiple-output analysis," **Journal of Health Economics,** 5, 107-127.

Koop, G. (1994). "Recent progress in applied Bayesian econometrics," **Journal of Economics Surveys,** 8, 1-34.

Koop, G., Osiewalski, J. and Steel, M.F.J. (1994). "Bayesian efficiency analysis with a flexible form: The AIM cost function," **Journal of Business and Economic Statistics,** 12, forthcoming.

Koop, G., Steel, M.F.J. and Osiewalski, J. (1994). "A Bayesian analysis of stochastic frontier models using Gibbs sampling," manuscript.

Kopp, R. and Diewert, W.E. (1982). "The decomposition of frontier cost function deviations into measures of technical and allocative inefficiency," **Journal of Econometrics**, 19, 319-332.

McCulloch, R. and Rossi, P.E. (1994). "An exact likelihood analysis of the multinomial probit model," **Journal of Econometrics**, forthcoming.

Meeusen, W. and van den Broeck, J. (1977). "Efficiency estimation from Cobb-Douglas production functions with composed error," **International Economic Review**, 8, 435-444.

Osiewalski, J. and Steel, M.F.J. (1994). "A Bayesian analysis of exogeneity in models pooling time-series and cross-section data," **Journal of Statistical Planning and Inference**, forthcoming.

Pitt, M. and Lee, L.-F. (1981). "The measurement and sources of technical inefficiency in the Indonesian weaving industry," **Journal of Development Economics**, 9, 43-64.

Schmidt, P. and Sickles, R. (1984). "Production frontiers and panel data," **Journal of Business and Economic Statistics**, 2, 367-374.

Simar, L. (1992). "Estimating efficiencies from frontier models with panel data: A comparison of parametric, non-parametric and semi-parametric methods with bootstrapping," **Journal of Productivity Analysis**, 3, 171-203.

Tierney, L. (1991), "Exploring posterior distributions using Markov chains," in **Computing Science and Statistics: Proceedings of the 23rd Symposium on the Interface**, E.M. Keramidas and S.M. Kaufman (eds.), Interface Foundation of North America.

Vita, M. (1990). "Exploring hospital production relationships with flexible functional forms," **Journal of Health Economics**, 9, 1-21.

Zellner, A. (1971). **An introduction to Bayesian inference in econometrics**, John Wiley and Sons: New York.

Data Appendix

The major data sets used in this analysis come from two independent sources: the American Hospital Association's (AHA) Annual Survey of Hospitals and the Health Care Financing Association (HCFA) Hospital Cost Reporting Information System (HCRIS) data files. Data were obtained for 1987-1991. The HCRIS files are cycles five through nine (1987-1991) of the Prospective Payment System (PPS). The sample represents all hospitals for which both AHA and PPS data were available, after eliminating specialty hospitals, all inclusive rate payers, and hospitals with fewer than 100 beds. The data bases of those 68 hospitals subject to all-payer systems of reimbursement were not comparable with those of the larger groups; the group of small hospitals exhibit cost structures that are distinctly different from those of hospitals having 100 or more beds.

To ensure a homogeneous panel we included only hospitals which produce all outputs and do not have a teaching mission (i.e. have neither a medical school affiliation nor a Council of Teaching Hospitals membership). As described in the text, we also excluded hospitals which had unusually high employee to patient ratios. The variables used for this procedure were the ratios of full-time equivalent clinical and non-clinical personnel to average daily census. Both came from the AHA survey.

The cost frontier incorporates the following. The dependent variable is the total of all direct costs, excluding Medicare non-reimbursable cost and capital related expenditures. The AHA measures of total discharges, inpatient days, facility bed total and outpatient visits are included as hospital outputs. To correct for case mix we use Medicare's DRG Case Mix Index as a fifth output category. The only available measure of input prices is that of wages. We use the indices of local area wage rates produced by HCFA for use in hospital cost research. Unfortunately, this has been rebased over time and, hence, the time series component of this series is questionable. To correct this problem, we divide each year's data by a constant to ensure that, on average, hospital wages are rising at the same rate as manufacturing wages. For capital stock we use the PPS calculation of total fixed assets.

In order to explain inefficiencies, we include measures of organization and market structure. A Herfindahl index is used as a measure of market competition, constructed using the county as market, and number of discharges as a measure of output from which to determine market shares. That the county is an acceptable alternative to a uniform geographic area in defining markets was found be Garnick et al. (1987). To see if different organizational structures affect efficiency, we include ownership dummies using the PPS categories of non-profit, profit and other (city, county or state facility).

Table 1: Description of Variables

C = costs (facility operating expenditure)

Y_1 = number of discharges
Y_2 = number of inpatient days
Y_3 = number of beds
Y_4 = number of outpatient visits
Y_5 = case mix index

P = wage index

K = capital stock

t = time trend

w_1 = intercept
w_2 = dummy variable for non-profit hospitals
w_3 = dummy variable for for-profit hospitals (dummy for government-run hospitals dropped)
w_4 = Herfindahl index averaged over time

Table 2: Prior Means of Efficiency for Different Values of r^*
(Prior standard deviations in brackets)

r^* =	.01	.20	.50	.80	.99
MIED CED	0.16 (0.26)	0.32 (0.32)	0.48 (0.34)	0.69 (0.30)	0.96 (0.11)
VED (m = 4)	0.10 (0.23)	0.19 (0.31)	0.28 (0.35)	0.42 (0.39)	0.78 (0.32)

Table 3: Prior Means of Efficiency for Different Values of a_j and g_j, $j>1$, $r^*=0.8$
(Prior standard deviations in brackets)

$a_j = g_j$ =	.6	1	5	10	25	100
VED (m = 4)	0.31 (0.39)	0.42 (0.39)	0.62 (0.34)	0.65 (0.32)	0.67 (0.31)	0.68 (0.31)

Table 4: Posterior Means of Elasticities for Minimum Cost Hospital
(Posterior standard deviations in brackets)

	P	Y_1	Y_2	Y_3	Y_4	Y_5	K
SIE	0.52 (0.11)	-0.07 (0.06)	0.42 (0.08)	0.30 (0.08)	-0.01 (0.02)	-0.26 (0.16)	0.03 (0.03)
MIED	1.19 (0.10)	0.09 (0.06)	0.47 (0.07)	0.32 (0.08)	-0.02 (0.02)	0.33 (0.15)	0.14 (0.03)
VED	1.18 (0.11)	0.09 (0.06)	0.48 (0.07)	0.32 (0.08)	-0.02 (0.02)	0.37 (0.15)	0.13 (0.03)
CED	1.22 (0.09)	0.09 (0.06)	0.47 (0.07)	0.33 (0.08)	-0.02 (0.02)	0.35 (0.15)	0.13 (0.03)
CE	1.18 (0.09)	0.16 (0.07)	0.56 (0.09)	0.20 (0.09)	-0.03 (0.02)	0.85 (0.18)	0.08 (0.03)

Table 5: Posterior Means of Elasticities for Median Cost Hospital
(Posterior standard deviations in brackets)

	P	Y_1	Y_2	Y_3	Y_4	Y_5	K
SIE	0.46 (0.09)	0.11 (0.04)	0.50 (0.05)	-0.11 (0.05)	0.05 (0.01)	0.44 (0.11)	0.07 (0.02)
MIED	0.81 (0.08)	0.30 (0.04)	0.53 (0.04)	-0.04 (0.05)	0.07 (0.01)	0.92 (0.11)	0.13 (0.02)
VED	0.78 (0.09)	0.30 (0.04)	0.53 (0.04)	-0.04 (0.05)	0.06 (0.02)	0.92 (0.10)	0.13 (0.02)
CED	0.82 (0.07)	0.30 (0.04)	0.53 (0.04)	-0.03 (0.05)	0.07 (0.02)	0.92 (0.10)	0.13 (0.02)
CE	0.77 (0.07)	0.36 (0.04)	0.46 (0.05)	-0.02 (0.05)	0.07 (0.02)	0.73 (0.11)	0.16 (0.02)

Table 6: Posterior Means of Elasticities for Maximum Cost Hospital
(Posterior standard deviations in brackets)

	P	Y_1	Y_2	Y_3	Y_4	Y_5	K
SIE	0.00 (0.11)	0.29 (0.06)	0.42 (0.07)	-0.07 (0.07)	0.05 (0.02)	-0.10 (0.13)	0.09 (0.03)
MIED	0.22 (0.09)	0.39 (0.06)	0.45 (0.07)	0.07 (0.07)	0.08 (0.02)	0.28 (0.12)	0.11 (0.03)
VED	0.20 (0.11)	0.40 (0.06)	0.44 (0.07)	0.07 (0.06)	0.08 (0.02)	0.26 (0.12)	0.11 (0.03)
CED	0.23 (0.09)	0.40 (0.06)	0.44 (0.07)	0.07 (0.06)	0.07 (0.02)	0.29 (0.12)	0.12 (0.03)
CE	0.46 (0.07)	0.41 (0.06)	0.38 (0.08)	-0.03 (0.07)	0.04 (0.02)	0.52 (0.14)	0.17 (0.02)

Table 7: Averages of Posterior Means of Efficiencies for Hospital Subgroups
(Averages of posterior standard deviations in brackets)

	Non-profit	For-profit	Govt.-run	all
SIE	0.46 (0.03)	0.48 (0.03)	0.50 (0.03)	0.47 (0.03)
MIED	0.85 (0.03)	0.79 (0.03)	0.86 (0.03)	0.84 (0.03)
VED	0.86 (0.03)	0.79 (0.03)	0.85 (0.03)	0.85 (0.03)
CED	0.86 (0.03)	0.80 (0.03)	0.87 (0.03)	0.85 (0.03)

Table 8: Posterior Means and Standard Deviations of γ

	γ_1	γ_2	γ_3	γ_4
Mean	2.30	-0.18	-0.14	-0.23
St. Dev.	0.69	0.50	0.14	0.27

FIGURE 1: MARGINAL PRIOR OF EFFICIENCIES

one-sided indiv. effects (r=0.8), SIE*

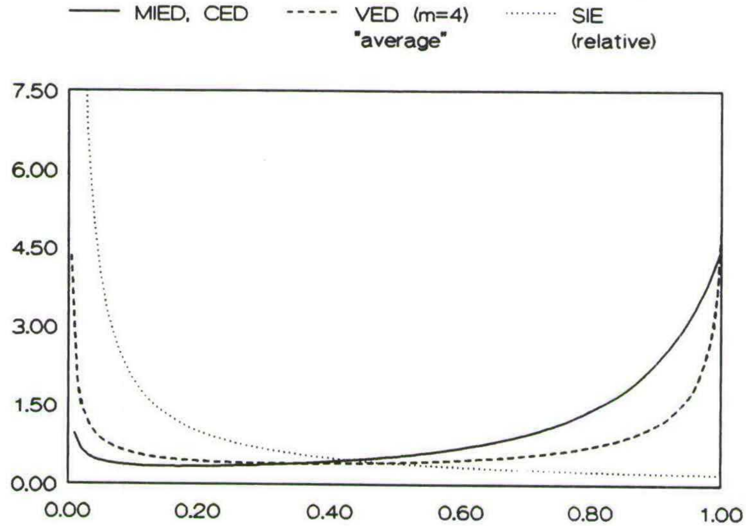


FIGURE 2: POSTERIORS FOR r(f)

$r*=0.8$

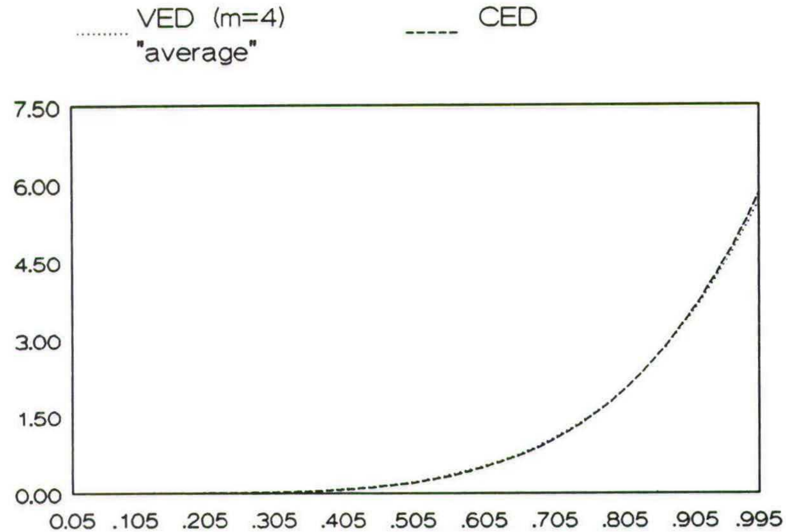


FIGURE 3: POSTERIORS FOR EFFICIENCIES

SIE model

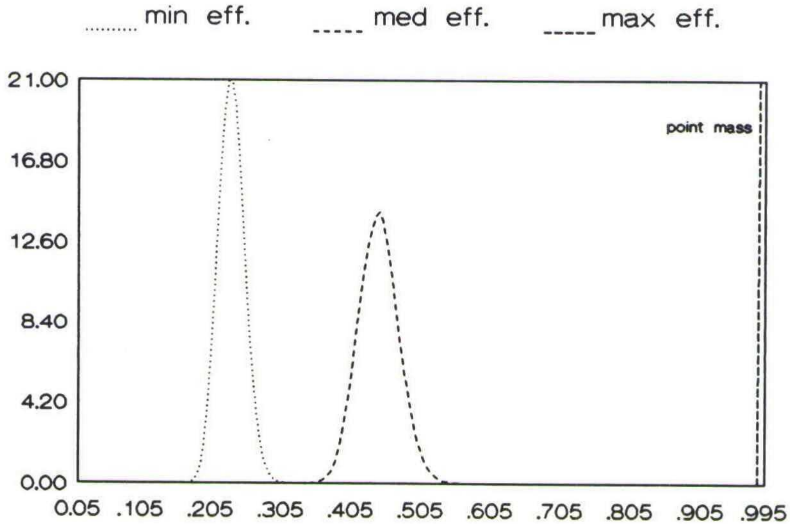


FIGURE 4: POSTERIORS FOR EFFICIENCIES

MIED model (r=0.8)*

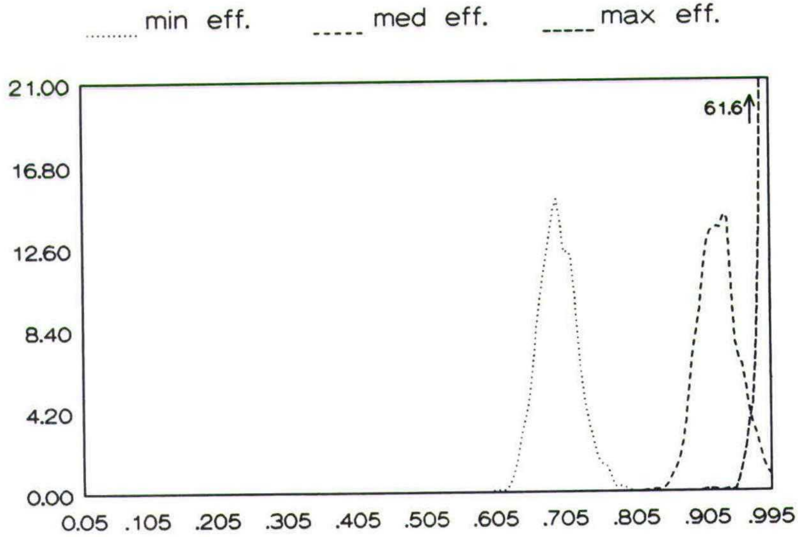


FIGURE 5: POSTERIORS FOR EFFICIENCIES

VED model (r=0.8, m=4)*

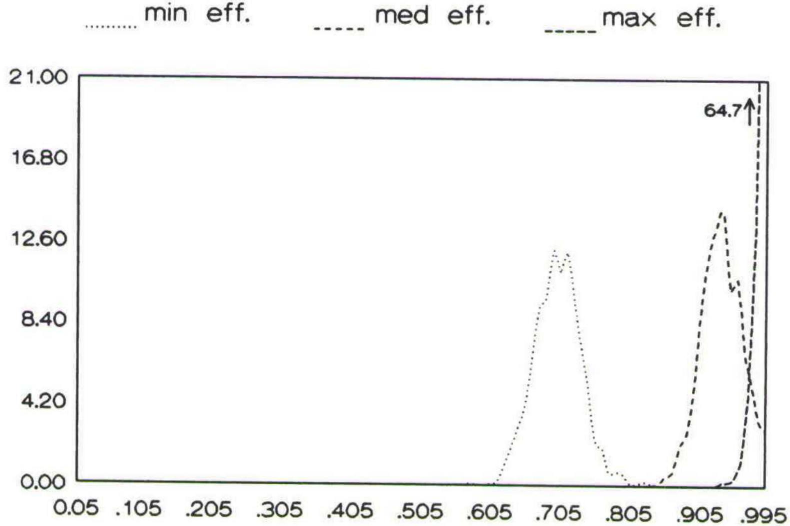


FIGURE 6: POSTERIORS FOR EFFICIENCIES

CED model (r=0.8)*

.... min eff. ----- med eff. ----- max eff.

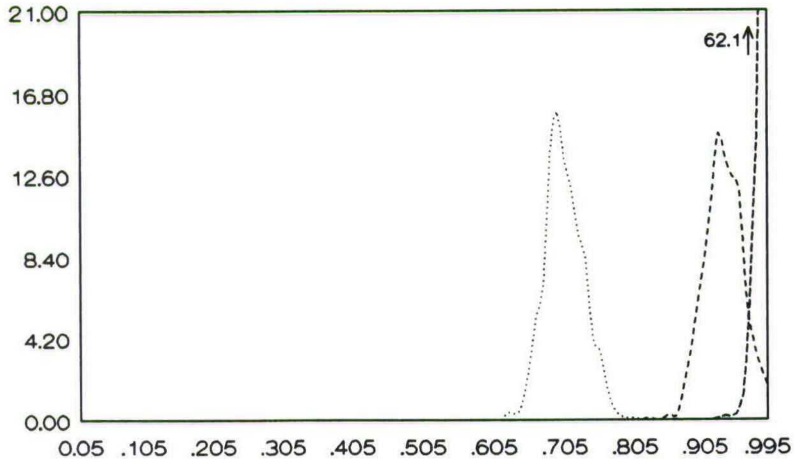


FIGURE 7: AVERAGE POSTERIOR MEAN EFF.

*as a function of r**

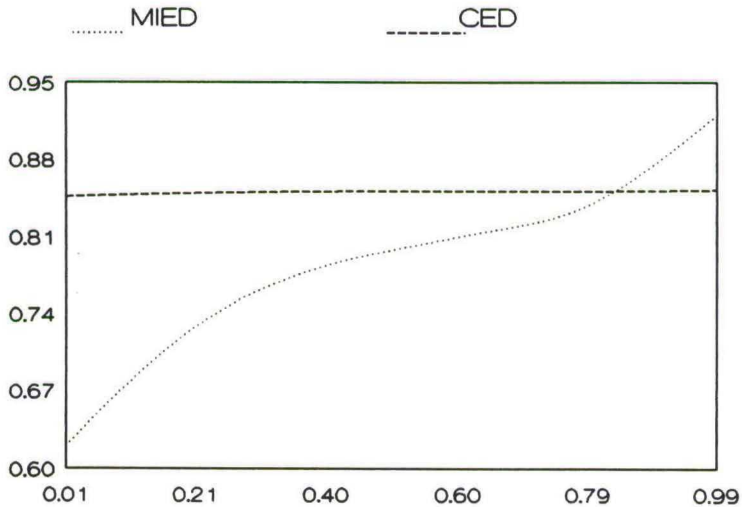


FIGURE 8: SENSITIVITY OF $p(r(f) \mid y, X)$

CED model

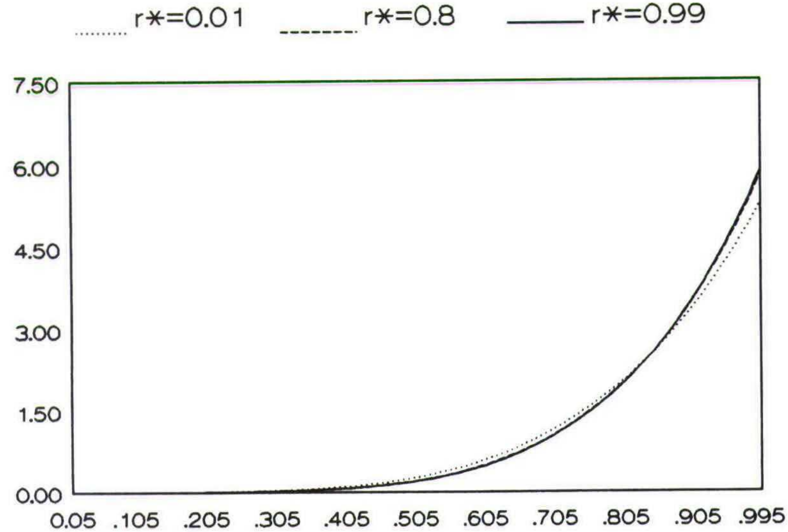
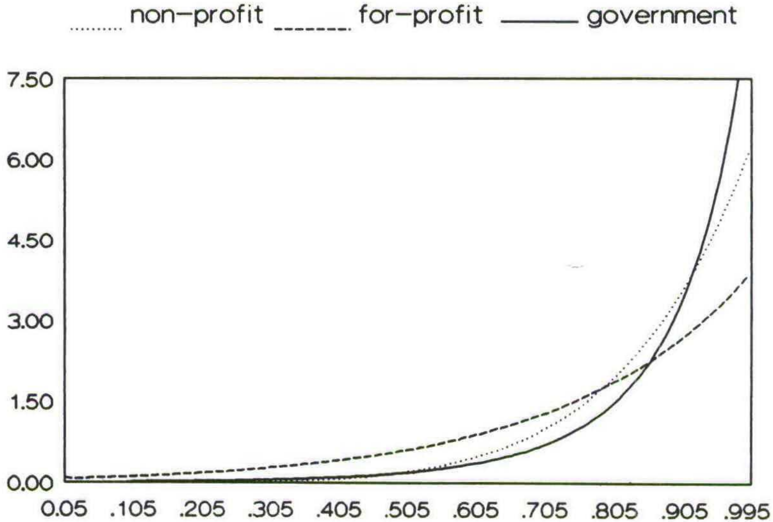


FIGURE 9: p(r(f) | y,X) FOR HOSPITAL TYPES

VED model (r=0.8, m=4)*



Discussion Paper Series, CentER, Tilburg University, The Netherlands:

(For previous papers please consult previous discussion papers.)

No.	Author(s)	Title
9340	T.C. To	Export Subsidies and Oligopoly with Switching Costs
9341	A. Demirgüç-Kunt and H. Huizinga	Barriers to Portfolio Investments in Emerging Stock Markets
9342	G.J. Almekinders	Theories on the Scope for Foreign Exchange Market Intervention
9343	E.R. van Dam and W.H. Haemers	Eigenvalues and the Diameter of Graphs
9344	H. Carlsson and S. Dasgupta	Noise-Proof Equilibria in Signaling Games
9345	F. van der Ploeg and A.L. Bovenberg	Environmental Policy, Public Goods and the Marginal Cost of Public Funds
9346	J.P.C. Blanc and R.D. van der Mei	The Power-series Algorithm Applied to Polling Systems with a Dormant Server
9347	J.P.C. Blanc	Performance Analysis and Optimization with the Power-series Algorithm
9348	R.M.W.J. Beetsma and F. van der Ploeg	Intramarginal Interventions, Bands and the Pattern of EMS Exchange Rate Distributions
9349	A. Simonovits	Intercohort Heterogeneity and Optimal Social Insurance Systems
9350	R.C. Douven and J.C. Engwerda	Is There Room for Convergence in the E.C.?
9351	F. Vella and M. Verbeek	Estimating and Interpreting Models with Endogenous Treatment Effects: The Relationship Between Competing Estimators of the Union Impact on Wages
9352	C. Meghir and G. Weber	Intertemporal Non-separability or Borrowing Restrictions? A Disaggregate Analysis Using the US CEX Panel
9353	V. Feltkamp	Alternative Axiomatic Characterizations of the Shapley and Banzhaf Values
9354	R.J. de Groof and M.A. van Tuijl	Aspects of Goods Market Integration. A Two-Country-Two -Sector Analysis
9355	Z. Yang	A Simplicial Algorithm for Computing Robust Stationary Points of a Continuous Function on the Unit Simplex

No.	Author(s)	Title
9356	E. van Damme and S. Hurkens	Commitment Robust Equilibria and Endogenous Timing
9357	W. Güth and B. Peleg	On Ring Formation In Auctions
9358	V. Bhaskar	Neutral Stability In Asymmetric Evolutionary Games
9359	F. Vella and M. Verbeek	Estimating and Testing Simultaneous Equation Panel Data Models with Censored Endogenous Variables
9360	W.B. van den Hout and J.P.C. Blanc	The Power-Series Algorithm Extended to the *BMAP/PH/*1 Queue
9361	R. Heuts and J. de Klein	An (s,q) Inventory Model with Stochastic and Interrelated Lead Times
9362	K.-E. Wärneryd	A Closer Look at Economic Psychology
9363	P.J.-J. Herings	On the Connectedness of the Set of Constrained Equilibria
9364	P.J.-J. Herings	A Note on "Macroeconomic Policy in a Two-Party System as a Repeated Game"
9365	F. van der Ploeg and A. L. Bovenberg	Direct Crowding Out, Optimal Taxation and Pollution Abatement
9366	M. Pradhan	Sector Participation in Labour Supply Models: Preferences or Rationing?
9367	H.G. Bloemen and A. Kapteyn	The Estimation of Utility Consistent Labor Supply Models by Means of Simulated Scores
9368	M.R. Baye, D. Kovenock and C.G. de Vries	The Solution to the Tullock Rent-Seeking Game When $R > 2$: Mixed-Strategy Equilibria and Mean Dissipation Rates
9369	T. van de Klundert and S. Smulders	The Welfare Consequences of Different Regimes of Oligopolistic Competition in a Growing Economy with Firm-Specific Knowledge
9370	G. van der Laan and D. Talman	Intersection Theorems on the Simplotope
9371	S. Muto	Alternating-Move Preplays and vN - M Stable Sets in Two Person Strategic Form Games
9372	S. Muto	Voters' Power in Indirect Voting Systems with Political Parties: the Square Root Effect
9373	S. Smulders and R. Gradus	Pollution Abatement and Long-term Growth
9374	C. Fernandez, J. Osiewalski and M.F.J. Steel	Marginal Equivalence in ν-Spherical Models

No.	Author(s)	Title
9375	E. van Damme	Evolutionary Game Theory
9376	P.M. Kort	Pollution Control and the Dynamics of the Firm: the Effects of Market Based Instruments on Optimal Firm Investments
9377	A. L. Bovenberg and F. van der Ploeg	Optimal Taxation, Public Goods and Environmental Policy with Involuntary Unemployment
9378	F. Thuijsman, B. Peleg, M. Amitai & A. Shmida	Automata, Matching and Foraging Behavior of Bees
9379	A. Lejour and H. Verbon	Capital Mobility and Social Insurance in an Integrated Market
9380	C. Fernandez, J. Osiewalski and M. Steel	The Continuous Multivariate Location-Scale Model Revisited: A Tale of Robustness
9381	F. de Jong	Specification, Solution and Estimation of a Discrete Time Target Zone Model of EMS Exchange Rates
9401	J.P.C. Kleijnen and R.Y. Rubinstein	Monte Carlo Sampling and Variance Reduction Techniques
9402	F.C. Drost and B.J.M. Werker	Closing the Garch Gap: Continuous Time Garch Modeling
9403	A. Kapteyn	The Measurement of Household Cost Functions: Revealed Preference Versus Subjective Measures
9404	H.G. Bloemen	Job Search, Search Intensity and Labour Market Transitions: An Empirical Exercise
9405	P.W.J. De Bijl	Moral Hazard and Noisy Information Disclosure
9406	A. De Waegenaere	Redistribution of Risk Through Incomplete Markets with Trading Constraints
9407	A. van den Nouweland, P. Borm, W. van Golstein Brouwers, R. Groot Bruinderink, and S. Tijs	A Game Theoretic Approach to Problems in Telecommunication
9408	A.L. Bovenberg and F. van der Ploeg	Consequences of Environmental Tax Reform for Involuntary Unemployment and Welfare
9409	P. Smit	Arnoldi Type Methods for Eigenvalue Calculation: Theory and Experiments
9410	J. Eichberger and D. Kelsey	Non-additive Beliefs and Game Theory
9411	N. Dagan, R. Serrano and O. Volij	A Non-cooperative View of Consistent Bankruptcy Rules

No.	Author(s)	Title
9412	H. Bester and E. Petrakis	Coupons and Oligopolistic Price Discrimination
9413	G. Koop, J. Osiewalski and M.F.J. Steel	Bayesian Efficiency Analysis with a Flexible Form: The AIM Cost Function
9414	C. Kilby	World Bank-Borrower Relations and Project Supervision
9415	H. Bester	A Bargaining Model of Financial Intermediation
9416	J.J.G. Lemmen and S.C.W. Eijffinger	The Price Approach to Financial Integration: Decomposing European Money Market Interest Rate Differentials
9417	J. de la Horra and C. Fernandez	Sensitivity to Prior Independence via Farlie-Gumbel -Morgenstern Model
9418	D. Talman and Z. Yang	A Simplicial Algorithm for Computing Proper Nash Equilibria of Finite Games
9419	H.J. Bierens	Nonparametric Cointegration Tests
9420	G. van der Laan, D. Talman and Z. Yang	Intersection Theorems on Polytopes
9421	R. van den Brink and R.P. Gilles	Ranking the Nodes in Directed and Weighted Directed Graphs
9422	A. van Soest	Youth Minimum Wage Rates: The Dutch Experience
9423	N. Dagan and O. Volij	Bilateral Comparisons and Consistent Fair Division Rules in the Context of Bankruptcy Problems
9424	R. van den Brink and P. Borm	Digraph Competitions and Cooperative Games
9425	P.H.M. Ruys and R.P. Gilles	The Interdependence between Production and Allocation
9426	T. Callan and A. van Soest	Family Labour Supply and Taxes in Ireland
9427	R.M.W.J. Beetsma and F. van der Ploeg	Macroeconomic Stabilisation and Intervention Policy under an Exchange Rate Band
9428	J.P.C. Kleijnen and W. van Groenendaal	Two-stage versus Sequential Sample-size Determination in Regression Analysis of Simulation Experiments
9429	M. Pradhan and A. van Soest	Household Labour Supply in Urban Areas of a Developing Country
9430	P.J.J. Herings	Endogenously Determined Price Rigidities
9431	H.A. Keuzenkamp and J.R. Magnus	On Tests and Significance in Econometrics

No.	Author(s)	Title
9432	C. Dang, D. Talman and Z. Wang	A Homotopy Approach to the Computation of Economic Equilibria on the Unit Simplex
9433	R. van den Brink	An Axiomatization of the Disjunctive Permission Value for Games with a Permission Structure
9434	C. Veld	Warrant Pricing: A Review of Empirical Research
9435	V. Feltkamp, S. Tijs and S. Muto	Bird's Tree Allocations Revisited
9436	G.-J. Otten, P. Borm, B. Peleg and S. Tijs	The MC-value for Monotonic NTU-Games
9437	S. Hurkens	Learning by Forgetful Players: From Primitive Formations to Persistent Retracts
9438	J.-J. Herings, D. Talman, and Z. Yang	The Computation of a Continuum of Constrained Equilibria
9439	E. Schaling and D. Smyth	The Effects of Inflation on Growth and Fluctuations in Dynamic Macroeconomic Models
9440	J. Arin and V. Feltkamp	The Nucleolus and Kernel of Veto-rich Transferable Utility Games
9441	P.-J. Jost	On the Role of Commitment in a Class of Signalling Problems
9442	J. Bendor, D. Mookherjee, and D. Ray	Aspirations, Adaptive Learning and Cooperation in Repeated Games
9443	G. van der Laan, D. Talman and Zaifu Yang	Modelling Cooperative Games in Permutational Structure
9444	G.J. Almekinders and S.C.W. Eijffinger	Accounting for Daily Bundesbank and Federal Reserve Intervention: A Friction Model with a GARCH Application
9445	A. De Waegenaere	Equilibria in Incomplete Financial Markets with Portfolio Constraints and Transaction Costs
9446	E. Schaling and D. Smyth	The Effects of Inflation on Growth and Fluctuations in Dynamic Macroeconomic Models
9447	G. Koop, J. Osiewalski and M.F.J. Steel	Hospital Efficiency Analysis Through Individual Effects: A Bayesian Approach

P.O. BOX 90153, 5000 LE TILBURG, THE NETHERLANDS